

# Raport PSI

Illia Dovhalenko, Maryia Babinskaya

June 2024

## 1 Wstęp

Projekt przedstawia model, który rozpoznaje 5 rodzajów zwierząt na zdjęciach.

## 2 Dane wejściowe

Dane wejściowe głównie są oparte na datasetcie z Kaggle (Animals classification dataset), który zawiera dane treningowe i walidacyjne 5 klas zwierząt. Ale ten zbiór nie zrównoważony, ponieważ zdjęcia psów i kotów zawierały tylko ich twarze, a zdjęcia innych zwierząt były różnego rodzaju. Dlatego postanowiliśmy rozszerzyć dataset o zdjęcia z innych zbiorów danych.

## 3 Strategia podziału danych wejściowych

Oryginalny zbiór danych zawierał już podział na dane treningowe i walidacyjne (15 tys. treningowych i 1.5 tys walidacyjnych). Po dodaniu dodatkowych zdjęć (39 tys treningowych i 3.3 tys walidacyjnych). Wykorzystuję batche o rozmiarze 128. Danymi testowymi występują losowe zdjęcia z internetu.

## 4 Problemy związane z projektem

Na początku mieliśmy około 15 tys. zdjęć treningowych, na których psy i koty miały tylko zdjęcia twarzy, a inne zwierzęta miały zdjęcia twarzy i ciała z różnych stron. Kiedy testowaliśmy model na zdjęciach kotów i psów z internetu, to on ich często rozpoznawał jako konie i słonie. Dlatego zwiększyliśmy zbiór treningowy zdjęciami ciała kotów i psów i też dodaliśmy dodatkowe zdjęcia lwów, słoni, koni. Teraz zbiór treningowy zawiera 39 tys. zdjęć, a walidacyjny 3.3 tys.

Po zwiększeniu zbioru pojawiły się problemy ze zdjęciami (do zwiększenia ich nie było): "Corrupt JPEG data: 214 extraneous bytes before marker 0xd9" i "Warning: unknown JFIF revision number 0.00". Próbowaliśmy usunąć te errorry poprzez funkcje, które usuwają problemowe zdjęcia, ale nie powiodło.

## 5 Uzasadnienie wyboru modelu

Sieci konwolucyjne są dobrze przystosowane do rozpoznawania obrazów. Możemy zmniejszać nasze obrazy do liczby naszych klas. Za pomocą operacji pooling (maxpooling) sieć może rozpoznać obiekt niezależnie od jego położenia. Łatwo można zapisać nasz model, i odtworzyć w innym notebooku.

## 6 Wyniki testowe i treningowe

W projekcie mamy dwa modele, jeden korzysta ze zwykłych zdjęć, a drugi ze zdjęć które mają rotację lub przybliżenie. Pierwszy model miał wysokie wyniki treningowe (na poziomie 90 proc.) ale źle działał w praktyce, dlatego korzystamy z drugiego modelu, które ma gorsze wyniki podczas treningu, ale lepiej działa w praktyce i ma około 75 proc. poprawnych odpowiedzi na losowych zdjęciach z internetu.

## 7 Analiza wyników

W ciągu 15 epok uczenia naszego modelu wynik walidacji ciągle rośnie od 0.57 w pierwszej epoce do 0.72 w 15. Testowanie przechodziło na zdjęciach z internetu. Otrzymaliśmy 72 procent poprawnie rozpoznanych zdjęć, ale część rozpoznał "prawie" poprawnie, ponieważ mógł rozpoznać psa jako kota (42.83 proc. że kot, ale 33.3 proc że pies) i dlatego warto patrzeć na procent przynależności do klasy, a nie tylko czy dobrze rozpoznał.

Nasz model ma też jawny problem z rudymi kotami, których on rozpoznaje za lwów. I z ciemnymi psami, których on rozpoznaje za konie.